Chapter VI
MEDICAL IMAGE REPRESENTATION-
LUNG-BREAST-CARDIAC IMAGES

6.1. LUNG CANCER AND TUMOR DETECTION USING DIRECTIONAL EDGE FILTERING TECHNIQUES

Lung cancer, breast cancer are the leading cause of cancer deaths in both men and women throughout the world. Breast cancers [81] in women over 45 years of age are due to following reasons: (1) Prior history of breast cancer in the family. It may be seen in mother or sister. (2) Nulliparity, that is women having no child or first child after 35 years of age be the reasons for the breast cancer. The other cancers are due to the use of tobacco and continuous smoking. While preventive methods are having some success, smokers indicate that there will be a long term need for improved techniques for early detection. The most common detection techniques known as radiography, cytologic analysis of sputum samples, fiber optic examination and computerized tomography (CT) scans are widely used. There are large numbers of diagnostic methods currently available, among which mammography is also the most reliable method.

The analysis of mammograms by computer [205] is roughly divided into three steps:

(i) Enhancement of pre-selected features
(ii) Localization of suspicious areas and
(iii) Classification of these areas into benign.

In order to detect the circumscribed masses, we define a suspicious area as area, brighter than its surrounding tissue, uniform density inside the area, an approximately circular shape of varying size and area having a fuzzy edge. Two examples of mammogram images are shown in fig.6.1.1 and fig.6.1.2.
Here the image enhancement by median filtering and edge enhancement by high-pass filtering and the edge detection is discussed.

**Filtering and edge detection**

**Median filtering:**

It has been found to remove noise from two-dimensional signals without blurring edges. The median filtering output is

$$X_{ij} = medican\{x_{r,s} : (r,s) \in N(i,j), (i,j) \in Z^2\} \quad \text{(6.1.1)}$$

Where $N(i,j)$ is the area in the image covered by the window $W(i,j)$. The window $W(i,j)$ is centered at image co-ordinates $(i,j)$ of a picture $\{x_{ij} : (i,j) \in Z^2\}$.

**High-Pass filtering**

The boundary information of the enhanced image is extracted for visual evaluation. A high-pass filter is used for this purpose. The high pass filter output is defined as:

$$g(x,y) = \left| \sum_{i} \sum_{j} f(x+i, y+j) - 9f(x, y) \right| \quad \text{(6.1.2)}$$
Directional filtering

The directional filtering could be expressed as

\[ G(x, y) = f(x, y) \otimes D \circ G(u, v, \theta(x, y)) \]

\[ = \sum_{u,v} f(x + u, y + v) \cdot D \circ G(u, v, \theta(x, y)) \]  ... (6.1.3)

Where \( D \circ G(\theta, x, y) = \Re(D \circ G(\theta, x, y)) = \Re(D \circ G(x, y)) \)

\[ = D \circ G(x \sin \theta - y \cos \theta, x \cos \theta + y \sin \theta) \]  ... (6.1.4)

\( f(x, y) \) denotes the intensity of original image and \( g(x, y) \) denotes the filtered image, \( \otimes \) represents this kind of non-conventional convolution. At each pixel, the operation between the two vectors is a typical inner product. DoG operator rotates based on the position of the centre pixel and guided by the rib norm orientation map of the posterior ribs. The following fig.6.1.4 shows the results after the performance of the directional filtering. We got the upper edge of the ribs and lower edge [235] of the ribs at the same time.

Fig.6.1.4.(a) The enhanced upper edge of ribs by directional filtering and (b) The enhanced lower edge of ribs by directional filtering.
**Edge detection**

An edge is defined as the boundary between two regions with relationally distinct gray level properties. Since the tumor is circular in shape, one alternative to detect tumors is to extract image edges and then look for ring like structures.

Different operators are used for edge detection such as Kirsch, Preewitt, Sobel, Quick, Frei, Chain and Canny [22]. From the observations, it is concluded that Canny operator gives more sharp and clear edges as compared to other operators.

![Fig. 6.1.5. Shows the edge detected images of fig.6.1.1(a) and 6.1.1(b) respectively.](image)

**Tumor detection using template matching**

For detecting a tumor, we use two characteristics that are approximately circular in shape and the brightness homogeneity of the tumor area.

One way to detect tumor is to find the image edges and then look for ring-like structures. However, in noisy or lightly textured images, a large number of noisy edges are extracted and edge tracking becomes difficult.

Hence, we use template matching which is based both on shape and brightness criteria. These criteria are defined by using the templates. The location of detected suspicious areas can then be obtained from the output of the matching operation, which can be visualized very clearly.
Fig. 6.1.6. Shows the output of template matching for fig 6.1.1(a) and 6.1.1(b) respectively.

Since the sizes of the tumors certainly differ from one another, templates of different sizes need to be used to obtain the exact size of the tumor.

**Similarity measure**

Suppose that ‘S’ is an image of size $L \times L$ array of pixels, each taking one of k gray level, and $W$ be the $M \times M$ template with $M << L$. Each $M \times M$ sub image of $S$ can be uniquely referenced by this upper left corner co-ordinates $(i,j)$.

The normalized cross-correlation measure is defined by

$$R(i, j) = \frac{\sum_{k=1}^{M} \sum_{m=1}^{M} \left[ (W(k, m) - \mu_w)^2 (S(i+k-1) - \mu_s(i, j))^2 \right]}{\sqrt{\sum_{k=1}^{M} \sum_{m=1}^{M} (W(k, m) - \mu_w)^2} \sum_{k=1}^{M} \sum_{m=1}^{M} (S(i+k-1) - \mu_s(i, j))^2} \quad \text{... (6.1.5)}$$

Where $\mu_w$ is the mean of the template and $\mu_s$ is the mean of the sub image centered at image point $(i,j)$. The template matching operation gives the output sub-image centered at that point. These values should be interpreted such that suspicious areas are rejected.
Effective criteria for selection of suspicious areas must be able to solve the following problems.

(i) Most suspicious areas should have the maximum cross-correlation value when being matched with different templates.

(ii) We do not have prior knowledge of the size and number of tumors in a mammogram film.

(iii) Some mammograms [74] have a rich image texture due to the presence of glands and fatty tissues leading to high cross-correlation values in the template matching stage. Some of these values may be even larger than those for some suspicious areas in other images.

**Percentile method**

We use a percentile method and classify a fixed percentage of locations as suspicious. The fixed percentage should be chosen so as to have no misses and a reasonably small number of false alarms. The template matching process only considers local information. The percentile method improves the template-matching step by taking into account the global image information.

**CAD Evaluation**

The aim of the modified computer aided design (CAD) system was only to detect the bright and round patches. The size of cancer varies from 7 mm to 20 mm. The suspicious regions are evaluated by CAD system specifically in the analysis of circularity and contrast.

This section presents various methods to detect the early stage of breast cancer, lung cancer and brain tumor. The first step towards tumor detection is image enhancement. The median filters are used.

The second step is edge enhancement for high pass filtering is used to enhance the edge and to make them sharp. The edge detection is also used for
tracking the edges. Several algorithms are used for edge detection. But, Canny edge detections [22] found to be more effective in obtaining clear and sharp edges.

The next step is concerned with tumor detection. Our method is based on template matching and is capable of detecting suspicious areas independent of their orientation and position. We use percentile method, which decreases the number of false alarms and also the non-suspicious areas to be considered as suspicious areas.

The lung cancer detection results using contrast and circularity assessment on temporal subtraction images showed a significant improvement over the current CAD method on the original chest images.

The results obtained in this method are quite encouraging. By combining the three criteria, namely, the contrast, the uniform density, and the circular shape of tumor areas, the detection algorithm is capable of detecting all tumor and cancer locations. The results for detecting tumors and locating its co-ordinates match with the radiologists’ results.

6.2. FUZZY INFERENCE SYSTEM APPLIED TO ENDOCARDIAL EDGE DETECTION

The goal of the edge detection process in a digital image is to determine the frontiers of all represented objects, based on automatic processing of the colour or gray level information in each present pixel. This procedure has many applications in image processing and computer vision. The two dimensional echo-cardiographic images [1, 36] have been used for the determination of cardiac performance. Various threshold-based methods have been applied either to the original ultrasound image or to the gradients of this image.

Fig.6.2.1 depicts the applied process, given as input image I. Two filters with impulse responses $h_{DX}$ and $h_{DY}$ are used to estimate derivatives in both
horizontal(x) and vertical(y) directions; points where filtered images $I_{DX}$ and $I_{DY}$ have values above the specified threshold are associated with vertical and horizontal edges, respectively. By computing norm-2, $\|x, y\| = \sqrt{x^2 + y^2}$ in each pixel of $I_{DX}$ and $I_{DY}$ and applying the threshold to the new output image, edges in arbitrary directions are detected.

![Diagram](image.png)

**Fig.6.2.1.** Edge detection system using LTI filters to estimate the image’s derivative in x($h_{DX}$) and y($h_{DY}$) directions.

The threshold value is to be estimated by root mean square value (RMS) associated to the input image [38]. Filters $h_{DX}$ and $h_{DY}$ are usually kernels with $5 \times 5$ or most often, $3 \times 3$ elements.

The most commonly used filter is so-called LoG (Laplacian-of-Gaussian). A recurring neural network with three outputs for each pixel of the original image is employed. Two of these outputs, $h$ and $v$, represent discontinuities in horizontal and vertical directions, respectively. The third output $f$, will be 1 or 0 after an iterative process, indicating if the corresponding pixel is or is not contained in an edge. An energy function is also defined and associated to all outputs of the network; the values of $h$, $v$ and $f$ are representations of the real edges of $I$. The neural network’s parameters are reduced to energy function value. This approach results in great computational effort. An image with $M \times N$ pixels, a neural network with $3 \times M \times N$ neurons is generated (for an image with $512 \times 512$ pixels, there will be 786432) [214].

**General description of the proposed system**

The different steps are involved in proposing the fuzzy inference system. We evaluate the efficiency of a FIS system applying it to edge detection problem. Sobel operators used a low-pass (mean) filter and a high-pass filter.
The key difference between this approach is that the gray level associate to pixel \((i, j)\) in the output image \(E\) depends not only on the pixel \((i, j)\) in each pre-processed image but also on some neighbour pixels. Each image \(DH\) and \(DV\) that results from applying Sobel operators is passed to the FIS system, the image composition is \(D = \sqrt{DH^2 + DV^2}\).

The developed fuzzy system’s purpose is to determine if pixel \((i, j)\) evaluated is or is not present in one of the image’s edges. This justifies the usage of a threshold in the binarization stage, when computing \(B(i, j)\).

![Figure 6.2.2. Fuzzy inference system (FIS) applied to edge (E) detection in image I.](image)

**Fig. 6.2.2.** Fuzzy inference system (FIS) applied to edge (E) detection in image I.

The different steps involved in fuzzy inference system are shown in fig. 6.2.3.

![Figure 6.2.3. General block diagram of proposed scheme.](image)

**Fig. 6.2.3.** General block diagram of proposed scheme.

**Implementation of the fuzzy inference system applied to edge detection**

Sobel operators \(h_{DH}\) and \(h_{DV}\) are kernels with \(3 \times 3\) elements given by

\[
h_{DH} = \begin{bmatrix}
-1 & 0 & 1 \\
-2 & 0 & 2 \\
-1 & 0 & 1
\end{bmatrix}
\]  \hspace{1cm} (6.2.1)

\[
h_{DV} = \begin{bmatrix}
1 & 2 & 1 \\
0 & 0 & 0 \\
-1 & -2 & -1
\end{bmatrix}
\]  \hspace{1cm} (6.2.3)
A 3 × 3 kernel is given by

\[
h_{HP} = \begin{bmatrix}
-1/16 & -1/8 & -1/16 \\
-1/8 & 3/4 & -1/8 \\
-1/16 & -1/8 & -1/16
\end{bmatrix}
\]  

...(6.2.3)

Fig.6.2.4(a) shows the magnitude of the frequency response of the filter described by equation (6.2.3). Notice that this magnitude is higher for higher frequencies, which means that the chosen filter presents the desired behaviour.

The MF is chosen in such a way that the gray level in each pixel of the output image is the arithmetic mean of the gray levels in a 5 × 5 neighbourhood of the same pixel in the input image

\[
H_{HP} = \frac{1}{25} \begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

...(6.2.4)

The magnitude of the frequency response of this filter is depicted in fig.6.2.4(b)

\[
DH = h_{DH} * 1 
\]

...(6.2.5)

\[
DV = h_{DV} * 1 
\]

...(6.2.6)

\[
HP = h_{HP} * 1 
\]

...(6.2.7)

\[
M = h_{MP} * 1 
\]

...(6.2.8)

**Fuzzy sets and fuzzy membership functions**

The input image I and the input variable M are associated to the linguistic variables “low”, “medium” and “high.” Gaussian functions are also adopted for the linguistic variables “low” and “medium”, but for the variable “high”, the input values will be restricted to the interval [0, 255]
Fuzzy logical operations and defuzzification method definitions

The functions adopted to implement the “and” (norm-T) and “or” (norm-S) operations are the minimum and maximum functions, respectively. The Mamdani method was chosen as the defuzzification procedure, which means that the fuzzy sets obtained by applying each inference rule to the input data are joined through the add function; the output of the system is then computed as the centroid of the resulting membership function.

Inference rules definitions

The fuzzy inference rules are defined in such a way that the FIS system output (“Edges”) is high only for those pixels belonging to edges in the image.

![Fig.6.2.4](image1)
Fig.6.2.4. (a) Discrete Fourier Transform of the high-pass and (b) The low-pass (mean) filter adopted in magnitude.

![Fig.6.2.5](image2)
Fig.6.2.5(a) Membership function of the fuzzy sets associated to the output and to the input M (b) The inputs DH, DV, HP.
A robustness to contrast and lighting variations are also in mind when these rules are established. The first three rules are

(i) (DH low) and (DV low) \rightarrow ("Edges" low).
(ii) (DH medium) and (DV medium) \rightarrow ("Edges" high)
(iii) (DH high) or (DV high) \rightarrow ("Edges" high).

To guarantee that edges in regions of relatively low contrast can be detected by using the two following rules

(iv) (DH medium) and (HP low) \rightarrow ("Edges" high).
(v) (DV medium) and (HP low) \rightarrow ("Edges" high).

Rules vi and vii are chosen in such a way as these regions are proportionally more affected by noise, but which does not favour edges by effect of noise.

(vi) (DV medium) and (M low) \rightarrow ("Edges" low).
(vii) (DH medium) and (M low) \rightarrow ("Edges" low).

Rules 8 to 11 are established to avoid forming double edges in the output image (they tend to appear due to shadows in the natural images). Considering that high variations in gray level in horizontal direction correspond to vertical edges, it is concluded that high values of $DH(i,j)$ and $DH(i,j \pm 1)$ do not imply edge pixels in $(i,j)$ and $(i,j \pm 1)$, simultaneously, the high values of $DV(i,j)$ and $DV(i \pm 1,j)$ do not correspond to edge pixels in $(i,j)$ and $(i \pm 1,j)$.

(viii) (DV high) and (DV(i+1,j) high) \rightarrow ("Edges" medium).
(ix) (DH high) and (DH(i,j+1) high) \rightarrow ("Edges" medium).
(x) (DV medium) and (DV(i+1,j) high) \rightarrow ("Edges" low).
(xi) (DH medium) and (DH(i,j+1) high) \rightarrow ("Edges" low).

Finally, rule 12 is defined to avoid including isolated pixels in the output image, favouring only continuous lines. It also avoids including points by effect of noise.

(xii) (DV(i,j+1) low) and (DH(i+1,j) low) and (DV(i,j-1) low) and (DH(i-1,j) low) \rightarrow ("Edges" low).
Materials and methods

Histogram equalization and median filtering has been incorporated over the echocardiogram images. The filter mask size used for median filtering here is $5 \times 5$. This method is purely subjective.

Fig. 6.2.6 (a) Input image (b) After Histogram Equalization (c) After Filtering.

So a fuzzy inference system is created with eight inputs each having 3 linguistic variables named “low”, ”medium”, ”high” and one output with two linguistic variable “edge”, ”non-edge” in both input and output side as shown in fig.6.2.6. Since the images are in the gray levels 0 and 255, the range of all FIS inputs should lie between $-255$ and 255 the output ranges between 0 and 255.

Fig. 6.2.7 (a) Membership functions of the fuzzy sets at input and (b) Output.
The function adopted to implement the “AND” and “OR” operations are the minimum and maximum functions respectively. The fuzzy inference rules are defined in such a way that the system output (“Edges”) is high only for those pixels belonging to edges in the input image. Totally 45 rules are written to guarantee the edges even in low contrast area of a noisy echo-cardiogram image [63, 65, 194].

Results

The different images are compared to that of the Sobel operator. The weight associated with each fuzzy rule is found to allow good results to be obtained while extracting edges of the image shown in fig.6.2.8. Where the contrast varies a lot from one region to another, during the performance tests, all parameters are kept constant.

By visual inspection we found that the Sobal and Perwitt operator fails to detect edges in low contrast region whereas in case of canny operator, the pixels not belonging to the edges also gets included. But the Fuzzy Inference System detects the edges and hence the endo-cardial borders are clearly shown even in low contrast region.
6.3. WEB-BASED DIAGNOSIS AND MEDICAL IMAGE RETRIEVAL SYSTEM

Medical diagnosis is supported by information provided by patient through interviews, general laboratory tests, and specialized examinations such as: mammography, magnetic resonance, and endoscopies. Radiological images represent a high percentage of these examinations and x-ray of thorax is one of the most frequently used.

Once the radiologist identifies a pathology, x-ray images are stored and scarcely used for further analysis. These images contain useful information to support either diagnosis of other patients or training of physicians in medical schools. In addition, their use is limited to accesses and availability. Therefore, storage solutions to facilitate cataloguing, browsing and retrieval of x-ray images are required.

Digital media overcomes the aforementioned shortcomings and disadvantages of physical storage such as deterioration and required room. In addition, it facilitates access and transmission of data.

In this section a system for effective storage and retrieval of medical images based on their content is presented.

This tool facilitates:

* Digital storage of radiological images and diagnostics.
* Management of related data, specifically clinic information.
* Web-based access.
* Queries by text and / or queries by image example.
* Data analysis through customized reports.

The system effectiveness is considered in three aspects: low access latency, accuracy, and scalability. Low access latency refers to the efficiency of the tool in terms of query/response delay. The success of any image retrieval system
depends on the speed the results are shown. Users are reluctant to accept search engines if they need to wait too long for the responses.

Accuracy describes the quality of the responses independently of the complexity of the content in the visual database. It does not make sense to develop efficient systems in terms of retrieval speed if the quality of the results is poor.

Scalability refers to the independence between the size of the image database and the query/response time delay. An approach is fully scalable if the time between query and response does not depend on the size of the database.

The implemented system combines visual features with symbolic descriptions for content-based radiological image retrieval. A method based on texture similarity applied on an interest zone. That is a sub-image containing a tumor, along with clinic data is used for cataloguing and searching x-ray images. Existing diagnostics are retrieved based on query by text and/or query by image example.

Finally, the system architecture uses an abstract. Application Program Interface (API) that provides a product-neutral interface between the front-end and database servers. It allows an application that is portable between servers from different manufactures.

**Theoretical background**

Digital images are usually obtained by converting continuous signals into digital and computer-readable binary format. Extraction and improvement of relevant information from these images are performed using digital image processing. It facilitates human interpretation and development of automatic visual information systems.
An image can be considered as a combination of different perceptual features. That is colour, texture, and shape; that can be used for searching and retrieving the image data. Content-based indexing and retrieval of images exploit automatic extraction of these features for managing information on large-scale image databases. In addition, a set of features can be combined with high level concepts. That is a diagnostic given by the radiologist, to generate interpretations of the image content.

**Digital image processing**

Digital image processing consists of five stages: acquisition, preprocessing, segmentation, representation and description, and recognizing and interpretation. The image acquisition stage pursues a representation of real-world objects through a digital image. It requires a sensor and a digitizer. The sensor is a physical device sensitive to certain band of the electromagnetic energy spectrum that produces an electric signal output proportional to the detected energy level. The digitizer is an electronic device which converts the electric signal from the sensor into a digital format [69, 73].

The preprocessing stage is addressed to improve the image quality to facilitate the next stages, in particular the segmentation stage. It is made either accentuating similarity or dissimilarity among pixels within a region or in different regions, respectively. Some representative preprocessing methods are: histogram modification [89], noise reduction [85], contrast increasing, and distortion correction. The segmentation stage consists of dividing the image into parts or objects to generate information at pixel level [39]. The unsupervised segmentation is one of the most difficult tasks in digital image processing.

The representation and description stage produces information at the highest abstraction level. The representation transforms sets of segmented pixels into more elaborated information elements such as contours and regions. Contours are focused on external features. That is, edge-based
representations [52]. Regions are concentrated on internal features. That is, texture-based representations. The description consists of feature selection and extraction to differentiate elements using quantitative information (image descriptors).

The recognizing and interpretation stage assigns labels based on information of its descriptors interpreting the meaning of a set of identified elements [10, 126].

**Content-Based Image Retrieval**

The use of low-level visual features to search and retrieve relevant information from image databases has drawn much attention in the recent years. As a result, several systems have been developed to search throughout image databases using visual primitives [49].

![Fig. 6.3.1. CBIR System Overview.](image)

Content-based retrieval systems (CBIR) use image descriptors to facilitate representation of features related to image content. A feature is a distinguishing primitive characteristic or attribute of an image, which is grouped into a feature vector. This vector can combine natural (luminance, texture, etc) and artificial (color histogram, etc.).

These descriptors can be linked at different levels of abstraction:
The raw data consists of elementary units together with some general attributes such as format, size, number of colours, etc. It is the lowest abstraction level.

Low-level content is characterized by features such as colour, shapes, textures, etc. It is the next higher abstraction level.

High-level semantic content contains high-level concepts such as objects and events. It is the highest abstraction level.

CBIR systems use several pattern recognition methods. In this context, a pattern recognition method is applied to feature extraction, clustering, indexing, and similarity-based image retrieval. In fig.6.3.1 an overview for a proposed content-based medical image retrieval is depicted.

System overview

The proposed system has two main parts: a control access sub-system and a content-based indexing and retrieval sub-system. Fig.6.3.2 illustrates the overall system architecture.

System architecture

The system architecture is based on web technologies and consists of three layers as shown in fig.6.3.3.

The first layer is accessed through a website. It performs validations and basic computing operations such as: image improvement and sub-image selection. This layer is supported by a Common Gateway Interface (CGI) to work on the server side and Applets to have part of the application functionalities on the Client side.

An extension of Servelets known as Java Server Pages (JSP) is applied to implement CGIs. JSP combines code in Java programming language and Hyper Text Markup Language (HTML). User's requests of JSP resources are
transformed by a web server into an object to facilitate further interactions during a session.

The second layer contains the application logic. It receives and processes the user's requests. It interacts with a repository through the third layer.

The third layer stores and retrieves information from a repository of persistent data. This layer uses Java Database Connectivity (JDBC) which provides a product-neutral interface between the front-end and database servers. It allows an application that is portable between servers from different manufactures.

**Content-based retrieval systems functionality**

The content-based image retrieval sub-system has three modules: image description module, storage module, and query/search module. The system functionality can be described as follows:
Let \( X = \{x_1, x_2, \ldots, x_N\} \) be an image \( f = \{f^0, \ldots, f^{N-1}\} \) be feature sets where \( f \) is a function of the image \( x \), and \( E' = \{e'_1, \ldots, e'_N\} \) be a pattern extracted from a feature vector \( f^0 \).

The input to the CBIR is a sub-image called interest zone which is extracted from a radiological image. In fig.6.3.4 a sample of a radiological image is shown. Fig.6.3.5 depicts a selected interest zone.

The image description module implements the procedure:

1. read \( x \)
2. extract \( f^0 \)
3. generate \( E' \)
4. read diagnosis

Feature extraction is applied on sub-images instead of the whole image. \( E' \) is based on the MPEG-7 Homogeneous Texture Descriptor, which provides a quantitative characterization of homogeneous texture regions for similarity-based image-to-image matching. It is based on computing the local spatial-frequency statistics of the texture.

The descriptor is computed by filtering the image with a bank of orientation and scale sensitive filters, and computing the mean and standard deviation of the filtered outputs in the frequency domain [130].

Semantics of Homogeneous Texture Descriptor (HTD) has 62 values and can be denoted as follows

\[
HTD = [f_{DC}, f_{SD}, e_1, e_2, \ldots, e_30, d_1, d_2, \ldots, d_{30}] \quad (6.3.1)
\]

Where, \( f_{DC} \) and \( f_{SD} \) are the mean and standard deviation of the remaining values. \( E_i \) and \( d_i \) are the nonlinearly scaled and quantized mean energy and energy deviation of the corresponding \( i^{th} \) channel.

Fig.6.3.6 presents the channel distribution. Table 6.3.1 shows the homogeneous texture descriptor for the sub-image depicted in fig.6.3.5.
Table 6.3.1: Homogeneous Texture Descriptor

<table>
<thead>
<tr>
<th>Average = 112</th>
<th>Standard Deviation = 101</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td></td>
</tr>
<tr>
<td>184</td>
<td>161</td>
</tr>
<tr>
<td>167</td>
<td>174</td>
</tr>
<tr>
<td>158</td>
<td>174</td>
</tr>
<tr>
<td>143</td>
<td>158</td>
</tr>
<tr>
<td>171</td>
<td>187</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Energy Deviation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>187</td>
<td>159</td>
</tr>
<tr>
<td>158</td>
<td>175</td>
</tr>
<tr>
<td>142</td>
<td>168</td>
</tr>
<tr>
<td>129</td>
<td>134</td>
</tr>
<tr>
<td>151</td>
<td>159</td>
</tr>
</tbody>
</table>

Fig.6.3.4. Lung cancer image sample.

Fig.6.3.5. Sub-image representing an interest zone.
The diagnosis corresponds to the semantic interpretation given by the radiologist to image x. It may include a suggested conduct or treatment to be followed.

Using strategies for semi-automatic image annotation, the semantic interpretation can be automatically propagated to other images [48].

The storage module implements the procedure:

1. **find** cluster for E'
2. **generate** index idx(x)
3. **link** idx(x) to clinic data
4. **store** x, E', and idx(x)

Given a user's request q the query/search module implements the procedure:

1. **read** q
2. **process** q
3. **search** content-based and clinic DB
4. **display** results

This procedure supports query by text and query by image example. Text-based predicates are used to extract semantic interpretations along with clinic data. An image is used to retrieve previous diagnostics of similar images using the corresponding descriptors and metrics.

The query result displays images and brief summaries of the associated information. If necessary, the links established by the storage module are used to extract detailed information.
The proposed system is robust and scalable. It retrieves and uses information contained in radiological imagery and diagnostics. The remote and on-line access facilitates contributions of expert radiologists and analysis based on previous cases.

This content-based diagnosis and medical image retrieval system is useful to support activities of far hospitals using a non-expensive infrastructure and internet access.

6.4. DETECTION OF BREAST CANCER

Breast cancer is one of the most common causes of death among women around the world [142]. Early detection of breast cancer is the best hope for reducing the death rate caused by the invasive disease. X-ray mammography is considered by general consent to be the best for early detection of breast cancer. Nevertheless, some problems stand up to this technology. First, it is a subjective task leading to multiple interpretations to each single mammogram. Second, mammograms have low contrast causing the analysis of these images to be difficult. Third, the number of mammograms to be analyzed by a radiologist per day is limited.

To overcome these problems, computer assisted diagnostic (CAD) techniques of breast cancer have been raised. We note that most of these techniques are performed in the spatial domain. Digital mammograms are huge in size, a feature inherent with medical images, leading to annoying latency in image processing and transmission and to wasted storage. This necessitates the use of compression techniques on image to reduce storage/transmission requirements. Many image compression techniques have been reported in literature [56]. The Digital Imaging and Communications in Medicine (DICOM) has adopted the Joint Photographic Expert Group 2000 (JPEG2000) standard to compress medical images [99].
In this section we present an algorithm for the detection of microcalcifications in the JPEG2000 domain. The algorithm is based on the statistical properties of the wavelet transform that the JPEG2000 coder employs. Simulation results have demonstrated that the proposed algorithm yields an excellent sensitivity and specificity at a reasonable false positive detection rate.

Related Work

The reported algorithms of breast cancer detection can be classified into two classes; spatial domain and wavelet based techniques.

(i) Spatial Domain Techniques

In these techniques, an image is preprocessed firstly in order to enhance its quality. Enhancement techniques involve morphological operations [146], Laplacian filtering [153], fuzzification [34], fractal analysis, higher order statistics [75], and histogram manipulation. The use of image enhancement increases the contrast of a mammogram image. Secondly, all regions that include micro calcifications are identified by segmenting the enhanced image using thresholding or neural network techniques. Thresholding techniques include histogram thresholding and entropy based thresholding. In this step, all candidate pixels being micro calcifications are identified. This might result in false positives. Lastly, false positive findings are eliminated using texture analysis.

The drawback of these methods is that digital mammograms are huge in size and it is expected that mammogram images will be compressed in order to reduce transmission / memory requirements. Therefore, it is necessary to decompress digital mammograms before applying the spatial domain techniques. Decompressing is a time consuming process and needs large memory size.
(ii) Wavelet based Techniques

In these techniques, a digital mammogram is first decomposed by a wavelet filter. The wavelet decomposition of a mammogram embeds the fine details in the high resolution levels of the decomposed image. The common wavelets that are used to detect micro calcifications in digital mammograms are Least Asymmetric Daubechies’ wavelet transform, with a finite basis of length [35, 239]. Secondly, the wavelet coefficients of the decomposed image are modified to enhance small details. Thirdly, a new enhanced image with clear micro calcifications and suppressed background is reconstructed. After that, thresholding is applied to the reconstructed image to detect micro calcifications as well as to eliminate false positives.

The drawbacks of these techniques are as follows: The time and space wasting still persist because mammogram images still need to be decompressed. Further complexity is added by the wavelet transformation step. Another disadvantage is in using wavelet filters that are not used widely in the medical imaging area.

We recall from previous section that many compression techniques have been reported in literature [56]. The JPEG2000 compression standard has been prominently useful in medical imaging as well as in image processing [237] and it is included in the Digital Imaging and Communications in Medicine (DICOM) [99]. This standard is introduced by the JPEG. We can find a detailed explanation of the JPEG2000 standard in [100].

Microclassification detection in the JPEG2000 Domain

We present the proposed algorithm for early detection of breast cancer in the domain of JPEG2000. The input to the algorithm is a compressed mammogram image. The output of the algorithm is a binary Image where its white spots point to micro calcifications that has been detected. The steps of the algorithm are illustrated in the block diagram that is shown in fig.6.4.1. In this
algorithm, the wavelet transform coefficients are extracted from the JPEG2000 mammogram image. The sub-images of the second and third level are extracted from the wavelet transformed image. The sub-images of the third level are down sampled and correlated with the second level sub-images. The correlated sub-images are scaled and then an image is resulted by applying interband standard deviation. Automatic thresholding is applied to yield candidates of micro calcifications. From those candidates, isolated pixels are eliminated and a final binary image is obtained. The following subsections explain in detail each step in the algorithm.

(i) Wavelet Transformed Image Extraction

In this step, the wavelet coefficient of the compressed image are extracted from the code stream of the JPEG2000 file. Fig.6.4.2 shows a digital mammogram in the spatial domain. The extracted coefficients of the wavelet transformed mammogram image that corresponds to the mammogram shown in fig.6.4.2. Fig.6.4.3 shows the wavelet resolution levels with the three oriented sub-images at each resolution level. Each resolution level of the wavelet transformed image includes four blocks. The upper left block is the approximation sub image that is supported to further decomposition. The remaining three blocks are the high-frequencies decomposed bands: horizontal (HL), Vertical (LH) and diagonal (HH) [117]. Our concern is the high frequency bands because micro calcifications are of high frequency. From our experiments, we found that the second and third levels of the wavelet transformed mammogram image are the best to apply our algorithms on. This is because although fine details are clear in
the higher levels, those levels are very noisy and resulted in high rates of false positives. On the other hand, fine details are lost in the lower levels and resulted in very low sensitivities.

![Mammogram Image](image1.png)

**Fig.6.4.2. Mammogram image in the spatial domain.**

![Wavelet Transformed Image](image2.png)

**Fig.6.4.3. Extracted wavelet transformed image.**

(ii) **Correlation**

Correlation is calculated by multiplying two adjacent scales of the wavelet transformed image in order to magnify significant structures and suppress noise. This is based on the fact that edge structures present observably at each sub band of the wavelet domain while noise decreases rapidly along the scales [249]. The aim of the correlation step here is not matching as in conventional correlation procedures. The matching between the two wavelet
scales exists since those scales are obtained from the same original image by correlating this original image with the translated and scaled mother wavelet function.

The correlation $C_{2,3}^i$ is performed simply by multiplying each wavelet coefficient in level 2, $W_2^i$ by its corresponding coefficient in the down sampled sub-images of level 3, $[W_3^i]_d$, $C_{2,3}^i = W_2^i[W_3^i]_d$ ...(6.4.1)

where $i = LH, HL, \text{and } HH$.

(iii) Scaling

Scaling is required to return the total power in the correlated coefficients $C_{2,3}^i$ to be the same as the total power in the wavelet coefficient set of the 2nd level, $W_2^i$, namely,

$$SC_{3,2}^i = C_{3,2}^i \times \frac{P_{w_2^i}}{P_{c_{3,2}^i}}$$ ...(6.4.2)

where

$$P_{w_2^i} = \sum_{x,y} W_2^i(x,y)W_2^i(x,y)$$ ...(6.4.3)

and

$$P_{c_{3,2}^i} = \sum_{x,y} C_{3,2}^i(x,y)C_{3,2}^i(x,y)$$ ...(6.4.4)

(iv) Determination of Standard Deviation

This step is required to identify pixels that are candidates to be micro calcifications. In this step, we calculate the standard deviation among the three corresponding wavelet coefficients in the three scaled sub-images. The output of this step is a sub-image called standard deviation (STD) where each pixel in it is resulted in calculating the standard deviation as defined in the following:

$$STD(x,y) = \sum_{i=HL, LH, HH} SC_{3,2}^i(x,y) - m$$ ...(6.4.5)

Where $m$ is the mean of the three oriented pixels and is given by,
The use of the standard deviation among bands that are resulted from levels 2 and 3 after correlating and scaling them is based on the fact that edges have high frequency components in the wavelet transformed image and have strong orientation in the wavelet domain bands where flat areas are suppressed because of its low frequency components and have no orientation. So micro calcifications that are considered as edges in the frequency domain will have the horizontal vertical and diagonal orientations in the bands and so the standard deviation among those will be much greater than the standard deviation among flat areas.

(v) Thresholding

Clusters of micro calcifications are isolated from all candidates that resulted from the previous step by obtaining a suitable threshold. The selection of the threshold is automatic and it is based on the non-Gaussian model of wavelets. That is, the highly peaked around zero [68] and the heavy tail distribution of the transform [8]. Briefly, we calculate the wavelet coefficients threshold by finding the valley in the histogram that resides after the median of the wavelet coefficients of the STD sub-image and assign it to our threshold. We choose the left most and the right most valleys. We start from the right half (i.e. the median) of the histogram, because the histogram of the micro calcifications affected sub-images are skewed right distribution (i.e. the heavy tail is located on the right side of the histogram) and so the micro calcifications are located in the right tail of the skewed right histogram. The left side of the histogram of the wavelet transformed mammogram image often contains wavelet coefficients for the black background. Next we move to the decision rule thresholding, namely,
Any pixel in the STD subimage that has a wavelet coefficient value which is larger than or equal to the threshold, TH, is set to one; else it is cleared to zero. The result of this step is a binary image f(x,y) contains all candidates of pixels to be micro calcifications. False signals are found here. To decrease the false positives, we calculate white dots in an overlapped 5 by 5 window that we move on the entire binary sub-image. If there is only one white point in the window we remove it else a cluster of micro calcifications is identified.

**Experimental results**

The Digital Database for Screening Mammography (DDSM) [43] is used to test the performance of the developed algorithms. The proposed algorithm is evaluated using 100 mammograms where 70 are affected by micro calcifications and 30 are normal. The diagnosed image and the specification file of each image are used to compare the detection results with the actual results. The mammogram images of the database are compressed using JPEG2000 to obtain 4 image sets. The first set is based on the lossless JPEG2000 compression (compression ratios 2:1 to 3:1). The second, third and fourth sets are based on the lossy JPEG2000 compression at compression ratios of 10:1, 50:1 and 100:1, respectively. Our simulations demonstrated setting the threshold as the minimum value of the least count wavelet coefficients set gives better results than setting it to the maximum value. Sensitivity and specificity are used to evaluate the proposed algorithm. Sensitivity evaluates how often the algorithm correctly identifies micro calcifications, while specificity evaluates how often the algorithm correctly specifies mammograms without micro calcifications.

The results of applying the proposed algorithm on the 4 sets of mammogram images are shown in table 6.4.1. For lossless compressed mammograms the proposed algorithm has yielded a sensitivity of 92% at a false positive detection rate of 4.7. For lossy compressed images at 10:1, 50:1, and
100:1, the obtained sensitivities are 92, 77 and 66 per cent respectively. The corresponding false positive rates are 2.4, 2.3, and 2.1 respectively. The specificity for the lossless compressed images is 87% at a false positive rate of 1.3. We note that the algorithm sensitivity and specificity are comparable to spatial domain and other wavelet based techniques. However, the proposed algorithm is based on standard image compression technique (JPEG2000) and has excellent false positive rates. Most importantly, the proposed technique saves time and memory space by eliminating the need to decompress mammograms for processing.

Table 6.4.1: Results of applying the proposed algorithm on DDSM Mammograms

<table>
<thead>
<tr>
<th>Compression Ratio</th>
<th>False Positive Rate</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>100:1</td>
<td>2.1</td>
<td>66%</td>
</tr>
<tr>
<td>50:1</td>
<td>2.3</td>
<td>77%</td>
</tr>
<tr>
<td>10.1</td>
<td>2.4</td>
<td>92%</td>
</tr>
<tr>
<td>Lossless</td>
<td>4.7</td>
<td>92%</td>
</tr>
</tbody>
</table>

6.5. EDGE DETECTION IN GATED CARDIAC NUCLEAR MEDICINE IMAGES

Gated cardiac nuclear medicine imaging uses an ECG trigger to gate the acquisition of images during the cardiac cycle. Multiple slices are usually obtained and they can be played back in a cine loop. This allows visualization of the wall motion of the ventricle as well as the calculation of the ejection fraction, which provides a very important evaluation of the ventricle function.

The edge detection software developed at ADAC Corporation for gated cardiac nuclear medicine images implements a radial search algorithm that assigns different threshold values for edges in different directions on the acquired images smoothed by a $5 \times 5$ spatial filter. These threshold values are stored in a memory table and they remain constant for all the slices under study.
Due to variations of brightness, contrast, and signal-to-noise ratio among different slices, this method may not be the most accurate method in extracting the left ventricle boundary. In this paper we propose an alternative approach using Mean Field Annealing (MFA) as a preprocessing tool for image smoothing and noise removal. The Laplacian operator is then used to extract edges from the MFA smoothing and noise removal. The Laplacian operator is then used to extract edges from the MFA smoothed images. Combined with the used input of initial boundary estimate, the extracted edge information is used in a minimum cost path search. The final boundary estimate is optimum with regard to boundary smoothness as well as the edge strength.

**Method**

A Markov Random Field based image processing technique, *Mean field annealing* (MFA), is used as a preprocessing tool for noise removal and image smoothing on the gated cardiac nuclear medicine images. A set of points plotted by a clinician have served as an initial boundary estimate. A second order operator has been applied to MFA preprocessed image data to get edge information. Both the initial boundary estimate and the edge detection information are then fed to a dynamic programming optimization algorithm, which minimized the cost function and also found the final boundary estimate.

**Mean field annealing smoothing**

MFA is a minimization strategy which is a deterministic approximation in Simulate Annealing. MFA has been applied to the restoration of locally-homogeneous and locally smooth range images [17, 18].

From the noise-corrupted measured image $g$, we want to estimate the unknown ideal image $f$. MFA restores the noisy image by maximizing the posteriori probability $p(f | g)$. Using Bayes’ rule, the a-posteriori probability can be expressed as the following proportionality:

$$
p(f | g) \propto p(g | f)p(f)$$  \hspace{1cm} \text{...}(6.5.1)
We model the noise as additive, independent, and stationary zero-mean Gaussian. Thus the noise term \( p(g \mid f) \) is of the form:

\[
p(g \mid f) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\sum_i (g_i - f_i)^2 / 2\sigma^2 \right) \quad \text{(6.5.2)}
\]

where the index \( i \) ranges all the pixels in the image. The noise term represents how the measured image resembles the “ideal” image.

The prior term, which represents our knowledge of the property of the ideal image, depends on \( f \) only, as is represented in the following form,

\[
p(f) = \exp \left( -\frac{\sum V_i}{T} \right) \quad \text{(6.5.3)}
\]

where \( V \) is a function of the neighbourhood pixels of pixel \( i \).

By taking the natural logarithm of both the left and the right side of equation (6.5.1) and changing the sign, the objective function is defined as:

\[
H(f, g) = H_n(f, g) + H_p(f) \quad \text{(6.5.4)}
\]

where \( H_n \) and \( H_p \) are the noise and prior terms respectively, and are chosen as follows:

\[
H_n(f) = \sum_i \left( \frac{f_i - g_i)^2}{2\sigma^2} \right) \quad \text{(6.5.5)}
\]

\[
H_p(f) = \sum_i \left( -\frac{b}{\sqrt{2\pi T}} \exp \left( -\frac{\Lambda_i^2}{2T^2} \right) \right) \quad \text{(6.5.6)}
\]

where \( \Lambda_i \) stands for some measure of the brightness change in the vicinity of pixel \( i \).

The restoration is performed to minimize \( H(f, g) \) because of the sign change. Minimizing \( H_n \) emphasizes that the restored image should resemble the
measured image. On the other hand, minimization of $H_p$ results in a restored image which satisfies our prior knowledge of the property of the ideal image.

In the gated cardiac nuclear medicine image, ventricle-ventricle and ventricle-atrium boundaries are roof edges. By choosing a linear-piecewise prior model, the roof edges of ventricle-ventricle and ventricle-atrium boundary are preserved. A second derivative operator, such as Laplacian, is then able to extract these edges. Here, we have a piecewise-linear model using quadratic variation was chosen as follows

$$
\Lambda^2 = \left( \frac{\partial^2 f}{\partial x^2} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 \quad \text{...(6.5.7)}
$$

This quadratic variation operator is implemented by convolving three kernels $(h_{xx}, h_{yy}, h_{xy})$ with the pixels in the 3 by 3 neighbourhood of pixels $i$, where

$$
h_{xx} = \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad h_{yy} = \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -2 & 0 \end{bmatrix}, \quad h_{xy} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{...(6.5.8)}
$$

and

$$
H(f, g) = \left( \frac{\sum (f_i - g_i)^2}{2\sigma^2} \right) - \beta \left( \frac{\sum \sum \frac{1}{T} \exp \left( -\frac{(f \otimes h)^2}{2T^2} \right) }{T} \right) \quad \text{...(6.5.9)}
$$

where $h$ is one of three kernels listed above.

The annealing was carried out iteratively by taking the gradient descent of (6.5.9), which results in the following,

$$
f_{k+1} = f_k - \gamma \frac{\partial}{\partial f_i} H(f, g) \quad \text{...(6.5.10)}
$$

and the resulting derivative works out to be

$$
\frac{\partial}{\partial f_i} H = \frac{(f_i - g_i)}{\sigma^2} + \frac{\beta}{T^3} \left( (f \otimes h)_i \exp \left( -\frac{(f \otimes h)_i^2}{2T^2} \right) \otimes h_{rev} \right) \quad \text{...(6.5.11)}
$$
where $\gamma$ is the relaxation ratio, and $h_{\text{rev}}$ is the reverse kernel of the convolution kernel (6.5.8).

**Boundary detection**

From the user input of the initial boundary estimate, a periodic, piecewise cubic spline was computed to interpolate these data points to form an initial boundary. At each user input point, the second derivative operator was applied to the pixels orthogonal to the initial estimated boundary on both sides. The number of pixels searched along the orthogonal line is variable based on the user selection. The obtained second derivatives generate an extracted data matrix, where the second derivatives along each orthogonal searching path form a column of this matrix, and the total number of columns is equal to the total number of the user input points.

The extracted data matrix, which contained both the user input estimate and the edge detection information is then used for the final boundary estimation.

To estimate a closed smooth boundary which tends to pass through the strongest edge points, it is proposed to minimize the following cost function [84, 232].

$$C(x) = (1 - \alpha) \sum_i (x_i - x_{i+1})^2 + \alpha \sum_i \{\max E_i - E_i(x_i)\} \quad \ldots(6.5.12)$$

where $\alpha$ is the weighing factor, $x_i$ indicates the row number of the boundary path in column $i$.

If two neighbouring points on the boundary pass through the same row $k$, that is $x_i = x_{i+1} = k$, then the first penalty term of (6.5.12) will be the minimum for those two-points. $E_i(x_i)$ (Second derivative) is the edge strength of the boundary point of column $i$, which is at column $i$ and row $x_i$. Max $E_i$ is the maximum edge strength of column $i$.

The weighing factor $\alpha$ can vary from 0 to 1. When $\alpha$ is set to 0, the smoothest path is found without using the edge information. On the other hand,
α=1 will result in the boundary passing through the strongest edge points with no penalty for sharp direction changes. For α between 0 and 1, the final boundary estimate will be a compromise between the boundary smoothness constraint and the edge strength.

**Result**

The MFA smoothing results are compared with the median filter and the 3 × 3 spatial filter. Fig.6.5.1(a) is the original blood pool image of the ventricle. The MFA restored image is shown in fig.6.5.1(b). Both the median filter and the 3 × 3 blurring filter are applied to the original image for smoothing. The results are shown in figs.6.5.1(c) and 6.5.1(d) respectively.

![Fig. 6.5.1(a)](image1) is the original noisy image. (b) shows the MFA smoothed image. (c) Median filtered original image. (d) 3 by 3 averaging kernel applied on the original image.

Simple smoothing operators such as the median filter operate on pixels within a local 3 × 3 or 5 × 5 neighbourhood and the same amount of smoothing
is performed throughout the whole image. As a result, the small features in a low variance area are blurred away and only the big features remain. MFA smoothing, on the other hand, preserves local features while removing random noise. The energy function (6.5.9) is minimized via MFA. We observe from fig. 6.5.1(b) the use of the $H_n$ of equation (6.5.2) has preserved the sharp noise-generated peaks as well as the roof edges. This is due to the assumption underlying equation (6.5.2); the noise is Gaussian. Since the noise in a nuclear medicine image is Poisson, a better noise model would be

$$p(g, f) = \frac{\left(\lambda f\right)^g \exp\left(-\lambda f\right)}{g!} \quad \ldots (6.5.13)$$

However, Han & Snyder has shown that equivalent results can be obtained by using local adaptive noise variance and local initial temperature in the annealing algorithm [76].

Fig.6.5.2(a) shows a cubic spline contour based on the initial user input estimate. The short lines which cross the boundary are the radial search lines. The second derivatives are calculated along these radial lines and they are used in the final estimate of the boundary. Fig.6.5.2(b) is the final estimated boundary. In comparison, the manually drawn contour (dotted line) is also shown in fig.6.5.2(b). Our result is very close to the contour drawn manually by
a radiologist. However, the semi-automatic generation of contours will be much more efficient than doing manually. For a multiple slice study, the boundary estimate from the first slice can be used to estimate the boundary in the following slices with no or minimum human intervention, which greatly reduces the heavy load of manual generation of contours on each slice.

The current implementation of the system utilizes a Gaussian noise model, and consequently does not find step edges with sufficient reliability to run fully automatically. Incorporating a more accurate Poisson model is currently under progress.